

Most frequent offence in Summer Months in San Francisco is LARCENY/THEFT with rate increase around mid August and around 7:30pm during the day.

1. Data preparation

```
sf <- read.csv("sanfrancisco_incidents_summer_2014.csv")

util.convert <- function(df) {
  df$Date <- as.Date(df$Date, '%m/%d/%Y')
  df$Descript <- as.character(df$Descript)
  df$Address <- as.character(df$Address)
  df$Location <- as.numeric(df$Location)

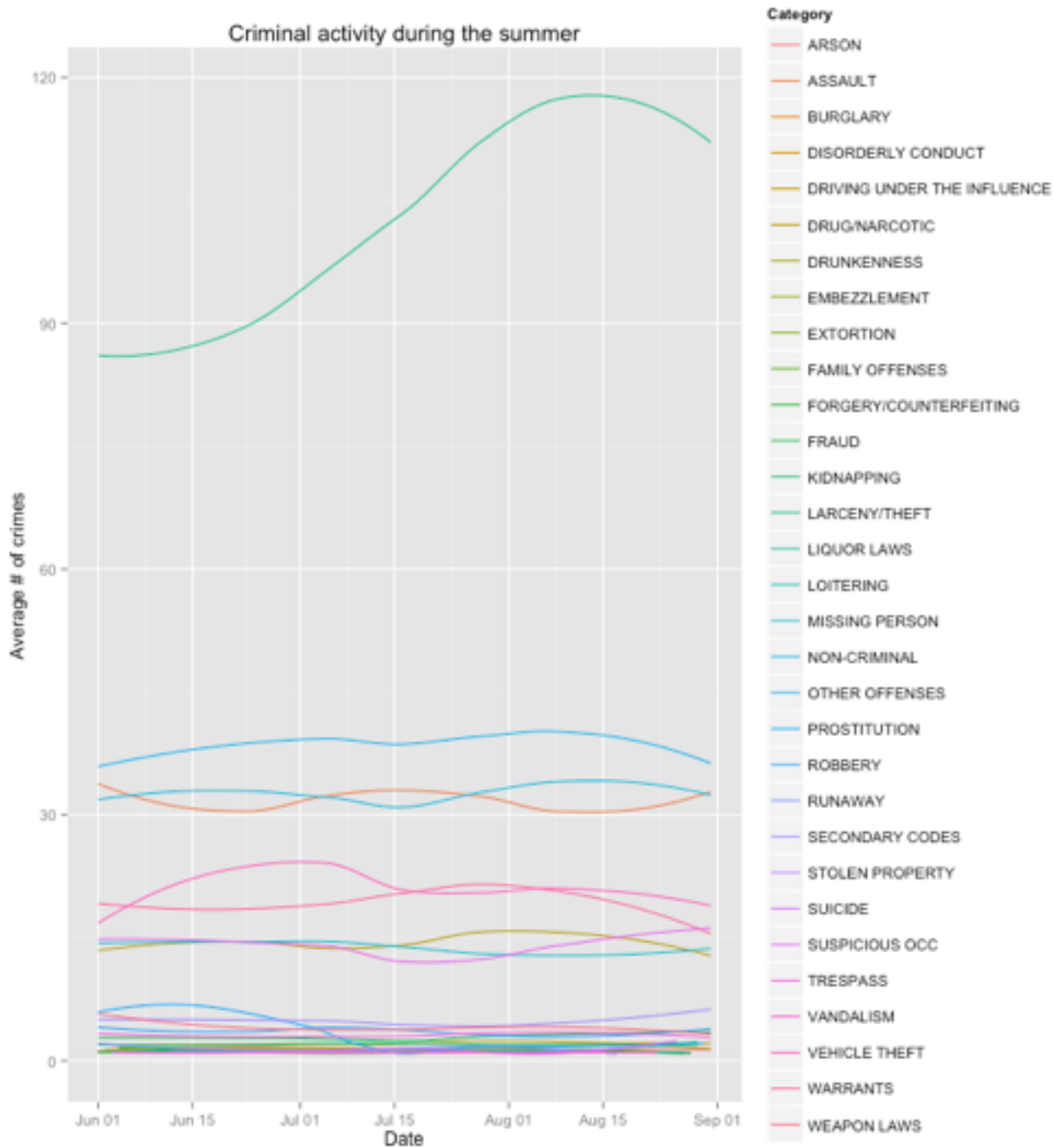
  return (df)
}

sf <- util.convert(sf)
sf_clean <- select(sf, Category, DayOfWeek, Date, Time,
PdDistrict, Resolution)
```

2. SF crime rate during the Summer

```
sf_grp_date <- group_by(sf_clean, Date, Category)
sf_grp_date <- summarise(sf_grp_date, c = n())
```

Please see results on the next page...



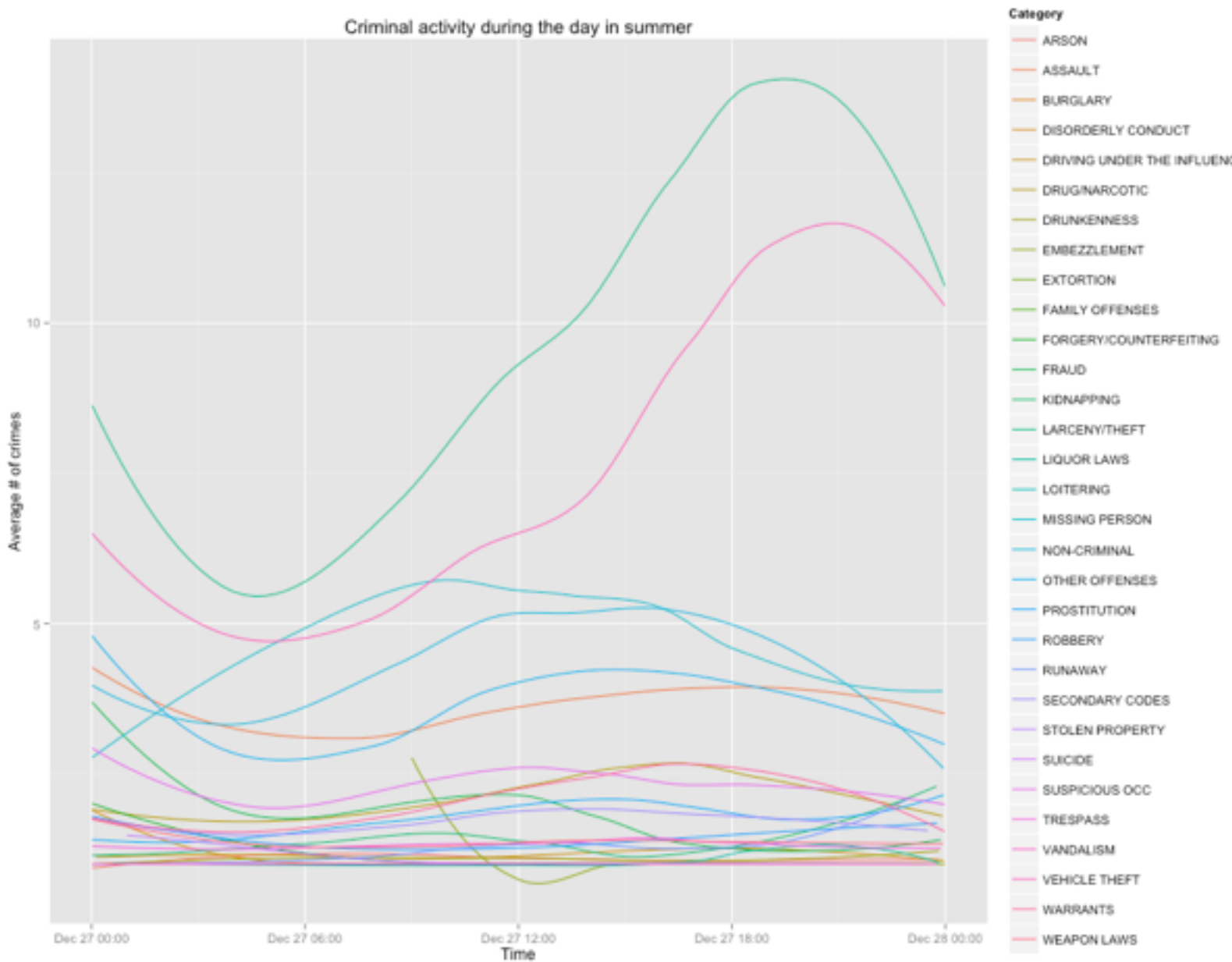
From the graphs above, we see that most frequent crime in SF during summer months is **LARCENY/THEFT** with peak around mid of August

3. SF crime rate during the day

```
sf_grp_time_cat <- group_by(sf_clean, Time, Category)
sf_grp_time_cat <- summarise(sf_grp_time_cat, c = n())
```

```
sf_grp_time_cat$Time <- strptime(sf_grp_time_cat$Time, format =
"%H:%M")
```

```
qplot(Time, c, data=sf_grp_time_cat, color=Category, geom =
c("smooth"), se = FALSE) + xlab("Time") + ylab("Average # of
crimes") + ggtitle("Criminal activity during the day in summer")
```



As shown above, **LARCENY/THEFT** and **VEHICLE THEFT** are the most frequent offences during summer day in SF with the peak around 7:30pm